IN THE CLAIMS

This listing of the claim will replace all prior versions and listings of claim in the present application.

Listing of Claims

Claims 1 and 2 (canceled).

3.   (currently amended)A similar document search method of searching a document similar to a seeds document from a document database, the method, comprising the steps of:

a first extracting step of extracting at least one characteristic word from the seeds document including desired retrieval contents;

a second extracting step of extracting as characteristic words of the seeds document, if the characteristic word extracted by the first extracting step is one of a compound characteristic phrase constructed by a plurality of constituent characteristic words, and the constituent characteristic words included in the compound characteristic phrase;

a step of calculating, according to the characteristic words extracted by the second extracting step, similarity between the seeds document and the document stored on the document database, by using the characteristic words including the compound characteristic phrase and the constituent characteristic words by which the compound characteristic phrase is constructed; and

a step of outputting a retrieval result as a result of the similarity calculated by the similarity calculating stepaccording to claim 1,

wherein the similarity calculating step includes the steps of:

2

a step of calculating a weighting coefficient corresponding to a distance

calculated by term appearance position on the seeds document between a

constituent characteristic word and another constituent characteristic word which are

extracted from one compound characteristic ~~word~~;phrase, and

a step of calculating similarity by multiplying the weighting coefficient.

Claims 4 and 5 (canceled).

6.   (currently amended)A similar document search system for searching a

document similar to a seeds document from a document database, the system,

comprising:

a document analyzer processor for extracting at least one characteristic word

from the seeds document including desired retrieval contents;

a characteristic word extractor processor for extracting as characteristic words

of the seeds document, if the characteristic word extracted by the document analyzer

processor is one of a compound characteristic phrase constructed by a plurality of

constituent characteristic words, and the constituent characteristic words included in

the compound characteristic phrase;

a seeds document similarity calculator processor for calculating, according to

the characteristic words extracted by the characteristic word extractor processor,

similarity between the seeds document and the document stored on the document

database, by using the characteristic words including the compound characteristic

3

phrase and the constituent characteristic words by which the compound

characteristic phrase is constructed; and

a retrieval result output processor for outputting a retrieval result as a result of

the similarity calculated by the seeds document similarity calculator

processor~~according to claim 4~~,

wherein the seeds document similarity calculator processor includes:

a weighting coefficient calculator processor for calculating a weighting

coefficient corresponding to a distance calculated by term appearance position on

the seeds document between a constituent characteristic word and another

constituent characteristic word which are extracted from one compound

characteristic ~~word;~~phrase, and

a calculator processor for calculating similarity by multiplying the weighting

coefficient.

Claims 7 and 8 (canceled).

9.   (currently amended)A program product for making a computer operate

as a similar document search system for searching a document similar to a seeds

document from a document database, the program comprising:

a document analyzer processor program for extracting at least one

characteristic word from the seeds document including desired retrieval contents;

a characteristic word extractor processor program for extracting as

characteristic words of the seeds document, if the characteristic word extracted by

the document analyzer processor program is one of a compound characteristic phrase constructed by a plurality of constituent characteristic words, and the constituent characteristic words included in the compound characteristic phrase;

a seeds document similarity calculator processor program for calculating, according to the characteristic words extracted by the characteristic word extractor processor program, similarity between the seeds document and the document stored on the document database, by using the characteristic words including the compound characteristic phrase and the constituent characteristic words by which the compound characteristic phrase is constructed; and

a retrieval result output processor program for outputting a retrieval result as a result of the similarity calculated by the seeds document similarity calculator processor program~~according to claim 7~~,

wherein the seeds document similarity calculator processor program includes:

a weighting coefficient calculator processor program for calculating a weighting coefficient corresponding to a distance calculated by term appearance position on the seeds document between a constituent characteristic word and another constituent characteristic word which are extracted from one compound characteristic ~~word;~~phrase, and

a calculator processor program for calculating similarity by multiplying the weighting coefficient.

5